



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/464,076	12/16/1999	BRIAN CRUICKSHANK	91436-209	7105
33000	7590	12/05/2003	EXAMINER	
			HAN, QI	
DOCKET CLERK		ART UNIT		PAPER NUMBER
P.O. DRAWER 800889		2654		
DALLAS, TX 75380		DATE MAILED: 12/05/2003		

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/464,076	CRUICKSHANK, BRIAN	
	Examiner Qi Han	Art Unit 2654	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1) Responsive to communication(s) filed on 26 September 2003.

2a) This action is FINAL. 2b) This action is non-final.

3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4) Claim(s) 1-12 and 14-20 is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5) Claim(s) _____ is/are allowed.

6) Claim(s) 1-12 and 14-20 is/are rejected.

7) Claim(s) _____ is/are objected to.

8) Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9) The specification is objected to by the Examiner.

10) The drawing(s) filed on _____ is/are: a) accepted or b) objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. §§ 119 and 120

12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a) All b) Some * c) None of:

1. Certified copies of the priority documents have been received.

2. Certified copies of the priority documents have been received in Application No. _____.

3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

13) Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application) since a specific reference was included in the first sentence of the specification or in an Application Data Sheet. 37 CFR 1.78.

a) The translation of the foreign language provisional application has been received.

14) Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121 since a specific reference was included in the first sentence of the specification or in an Application Data Sheet. 37 CFR 1.78.

Attachment(s)

1) Notice of References Cited (PTO-892) 4) Interview Summary (PTO-413) Paper No(s). _____ .

2) Notice of Draftsperson's Patent Drawing Review (PTO-948) 5) Notice of Informal Patent Application (PTO-152)

3) Information Disclosure Statement(s) (PTO-1449) Paper No(s) _____ . 6) Other: _____ .

DETAILED ACTION

1. The text of those sections of Title 35, U.S. Code not included in this action can be found in a prior Office action.

Response to Amendment

2. This communication is responsive to the applicant's amendment dated 09/26/2003 (Paper 12). Applicant amended claims 1, 9-12 and 14, and added new claims 19-20.

Response to Arguments

3. Applicant's arguments filed on 09/26/2003 (Paper 12) have been fully considered.

In response the applicant's argument (regarding rejection under 35 USC 102) that "the Office Action fails to show that Sharman generates a list of words, prefixes and/or suffixes using the dictionary" (paper 12, page 12,), examiner respectfully disagrees with applicant, because Sharman did disclose and/or suggest all the elements of claimed including a listing (list) of words, prefixes and/or suffixes (column 5, lines 18-40, also see the claim rejection in new office action)

In response to applicant's argument about obviousness for combining prior art (paper 12, page 14, paragraph 1), the examiner recognizes that obviousness can only be established by combining or modifying the teachings of the prior art to produce the claimed invention where

there is some teaching, suggestion, or motivation to do so found either in the references themselves or in the knowledge generally available to one of ordinary skill in the art. See *In re Fine*, 837 F.2d 1071, 5 USPQ2d 1596 (Fed. Cir. 1988) and *In re Jones*, 958 F.2d 347, 21 USPQ2d 1941 (Fed. Cir. 1992). In this case, obviousness is based on prior art teaching for claims 1 and 9-12 (see detail in the claim rejection of new office action); and based on common knowledge in the art for claims 7 (see detail in the claim rejection of new office action) and 8 because using a symbol, such as "\$", for distinguishing a data from other data, is common sense in the art for programming.

Regarding the argument for claim 1 and 9-12 (paper 12, page 14, paragraph 3 through page 15, paragraph 2) under rejection under 35 USC 103, the response is base on the same reason as rejection under 35 USC 102 (see above) because applicant argued same issue.

Regarding the argument for claim 14 (paper 12, page 14, paragraph 3), examiner rewrites the rejection (see detail in the claim rejection of new office action).

Claim Rejections - 35 USC § 101

35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

4. Claim 14-15 and 19-20 are rejected under 35 U.S.C. 101 because the claimed invention lacks patentable utility.

Claims 14-15 and 19-20 recite a data structure having several fields without any functional relationship to any hardware and/or software functionality, so that it is non-functional, therefore it lacks patentable utility. Even though each field in the data structure are described with different intention for use, such as “a field for a textual unit”, it does not change the nature that the claimed data structure is merely the template with several fields that any data can be applied to it and no functional utility.

5. To expedite a complete examination of the instant application the claims rejection under 35 U.S.C 101 (nonstatutory) above are further rejected as set forth below in anticipation of applicant amending these claims to place them within the four statutory categories of invention.

Claim Rejections - 35 USC § 102

6. Claim 6 is rejected under 35 U.S.C. 102(b) as being anticipated by Sharman.

Regarding **claim 6**, Sharman discloses a text to speech system. Sharman further discloses a linguistic processor for various linguistic processes comprising:

receiving a text file, (column 2, line 2-3, ‘input text’; column 5, lines 1-2, ‘obtain input from a source, such as ... a stored file’);

parsing said text file into textual units, where each said parsed textual unit is one of a word, a prefix or a suffix, (column 5, lines 3-40, ‘split input text into tokens (**words**)’, implement special rules ‘to map lexical items into canonical **word form**’, ‘using a dictionary look-up’, ‘remove any possible **prefix** or **suffix**’); and

for each one of said parsed textual units, if said one of said parsed textual units corresponds to a stored textual unit in a vocabulary of textual units, adding said stored textual unit to a list, (column 2, lines 1-2, 'generating a listing (list) of speech segments (equivalent textual units) ... from the input text', herein the list is inherently stored in a buffer; column 5, lines 26-27 'to see if the word is related to one that is already in the dictionary'; column 6, lines 61-66 and column 7, Table 1, 'output unit represents the size of the **text unit** (including word, phoneme)' used for different process stages; column 7, lines 45-66, '**output buffer** is also used when a component produces several outputs units for each input unit that it receives', herein inherently including adding prefix and suffix to the buffer because without storing them in the buffer the system cannot output required speech).

Claim Rejections - 35 USC § 103

7. Claims 1-5, 9-12 and 18 are rejected under 35 U.S.C. 103(a) as being unpatentable over Sharman (USPN 5,774,854) in view of Hata et al. (USPN 5,878,393) hereinafter referenced as Hata.

Regarding **claim 1**, Sharman discloses a text to speech system, comprises: receiving a list of textual units, (column 2, lines 1-2, 'a linguistic processor for generating a listing of speech segments (equivalent textual units) ... from the input text'), where said textual units in the list comprise words, prefixes and suffixes, (column 5, lines 18-27, 'removing any possible **prefix or suffix**, to see if the **word**, is related to one that is already in the dictionary'); acoustic processor 220 (Figs. 2 and 4) preparing acoustic data by using diphone library 420 (Fig.4) and diphone concatenation unit 415 (PSOLA) (column 6, lines 25-38), storing the

result in the output buffer (column 6, lines 23-24), wherein the output buffer is used when a component produces **several output units** for each input unit that receives (column 7, lines 54-67), including token, word, phoneme, syllable (herein inherently including prefix and suffix because they must be stored in the buffer in order to obtain an output speech) (column 7, table 1); and producing the acoustic waveform (output signal) (column 6, lines 55-60); which corresponds to the claimed “for each textual unit in the list locating an associated speech” “in memory and appending said associated speech” “to an output signal”

But, Sharman fails to explicitly disclose utilizing “speech sample” for the speech in the diphone library for the phonetic data, though he recites that a diphone library 420 (Fig. 4) effectively contains prerecorded segments of diphones (column 6, line 25). However, the examiner contends that the concept of providing speech sample as phonetic data was well known, as taught by Hata.

In the same field of endeavor, Hata discloses a high quality concatenative reading system. Hata further discloses that the system has a dictionary of sampled sounds 40 (Fig. 1) (column 3, 42-45) and the individual speech samples represent discrete units of speech, such as phonemes or words (column 3, line 26-31). Furthermore, Hata discloses multiple buffers for storing text and speech data in different processing stages, including input buffer 44 (Fig. 1), word list buffer 48, and sample list buffer 54 (column 5, lines 6-26), which is inherently capable of storing the textural list as the claimed.

Therefore, it would have been obvious to one of ordinary skill in the art at time the invention was made to combine Sharman and Hata, to specifically provide stored speech sample

for generating sound data, as taught by Hata, for the purpose of producing high-quality concatenated synthesized speech (Hata: column 2, line 57).

Regarding **claim 2**, Sharman and Hata disclose everything claimed, as applied above (see claim 1). Sharman further suggests that: (i) at substring level, it is useful to include some back-up mechanism to be able to process words that are not in the dictionary (column 5, line 24); (ii) at phoneme level, it is again using a dictionary look-up table, augmented with general purpose rules for words not in the dictionary (column 5, line 34); which is equivalent to use “secondary text-to-speech engine”. Further more, Sharman discloses that the phoneme data and other portion of data are sent to acoustic processor to produce output data stored in the output buffer 590 (Fig. 5) (column 8, line 23-24). This corresponds to the claimed “wherein one said textual unit in said list is indicated as not having an associated speech sample in memory and said method further comprises: passing said indicated textual unit to a secondary text to speech engine; receiving a speech sample converted from said indicated textual unit from said secondary text to speech engine; and appending said converted speech sample to said output signal.”

Regarding **claim 3**, Sharman and Hata disclose everything claimed, as applied above (see claim 2). But, Sharman fails to explicitly disclose that “said secondary text-to-speech engine comprises a phonetic text-to-speech engine based on a **voice talent**”. However, the examiner contends that the concept of utilizing a phonetic text-to-speech engine based on stored and processed speech sample (herein equivalent to a voice talent) was well known, as taught by Hata.

Hata further discloses that the system has a dictionary of sampled sounds 40 (Fig. 1) (column 3, 42-45) and the individual speech samples (equivalent to voice talent) each represent discrete units of speech, such as phonemes or words (column 3, line 26-31)

Therefore, it would have been obvious to one of ordinary skill in the art at time the invention was made to modify Sharman by specifically providing a phonetic text-to-speech engine based on stored and processed speech sample for a TTS engine, as taught by Hata, for the purpose of increasing sound quality for the system.

Regarding **claim 4**, Sharman and Hata disclose everything claimed, as applied above (see claim 1). Sharman in view of Hata discloses that processing input text at the substring level is based on a syllabified word (Sharman: column 5, line 31), so that combined system inherently satisfies all limitation elements as claimed “wherein a consecutive plurality of said textual units in said list represent a whole word, said method further comprising: for each textual unit in said consecutive plurality of said textual units, locating an associated speech sample in said memory; creating a speech unit by splicing together said plurality of associated speech samples; and appending said speech unit to said output signal.”

Regarding **claim 5**, Sharman and Hata disclose everything claimed, as applied above (see claim 4). Sharman further discloses components of identifying diphones 410 (Fig. 4), diphone library 420 and diphone concatenation 415 for overcoming audible discontinuities (column 6, lines 34-40), which corresponds to the claimed “after said splicing, processing said speech unit to remove discontinuities.”

Regarding **claim 9**, it discloses an apparatus, which corresponds to the method of claim 1; the apparatus is obvious in that it simply provides structure for the functionality found in claim 1.

Regarding **claim 10**, it discloses an apparatus, which corresponds to the method of claim 1; the apparatus is obvious in that it simply provides structure for the functionality

found in claim 1. In addition, Sharman specifically discloses that the TTS system includes two microprocessors (column 3, line 17), which corresponds to the claimed “a text to speech converter comprising a processor operable to . . .”.

Regarding **claim 11**, it discloses an apparatus, which corresponds to the method of claim 1; the apparatus is obvious in that it simply provides structure for the functionality found in claim 1. In addition, Sharman specifically discloses that an arrangement is particularly suitable for a workstation (equivalent to computer) equipped with an adapter card with its own DSP (equivalent to processor) (column 3, line 21), which corresponds to the claimed “a computer readable medium for providing program control to a processor, said processor included in a text to speech converter, said computer readable medium adapting said processor to be operable to . . .”.

Regarding **claim 12**, it discloses an apparatus, which corresponds to a combination of the method of claim 1 and the method of claim 6; the apparatus is obvious in that it simply provides structure for the functionality found in claim 1 and claim 6.

Regarding **claim 13**, it is canceled.

Regarding **claim 18**, it depends on the claim 12; and it discloses an apparatus, which corresponds to a combination of the method of claim 7 and the method of claim 16; the apparatus is obvious in that it simply provides structure for the functionality found in claim 7 and claim 16.

8. Claims 7 and 16 are rejected under 35 U.S.C. 103(a) as being unpatentable over Sharman in view of Microsoft Press (“Computer Dictionary”, page 298) hereinafter referenced as R1.

Regarding claim 7, Sharman discloses everything claimed, as applied above (see claim 6). Sharman particularly discloses that apart from using a dictionary look-up, “it is useful to include some back-up mechanism to be able to process words that are not in the dictionary” (column 5, lines 24-26), which is corresponding to the claimed “if said one of said parsed textual units does not correspond to one of said stored textual units” and “as being out of vocabulary.” Sharman further recites that “the output unit represents the size of the text unit (e.g. word, sentence, phoneme); for many stages this is accompanied by additional information for that unit (e.g., duration, part of speech etc.)” (column 6, line 59 to column 7, line 2), which suggests that the text unit may be different in each of processing stages. But, Sharman fails to explicitly disclose to mark a text unit that does not match the one either in dictionary or by rule sets. However, the examiner contends that the concept of marking a text unit data was well known, as taught by R1.

R1 is a popular computer dictionary that gives common meaning and explanation of words or phrases in computer related arts. R1 further discloses that one of the common meanings of the word “mark” is “in applications and data storage, a symbol or other device used to distinguish one item from others like it” (page 298, entry “mark”), so that when using “mark” as a verb, it can be interpreted as an action to mark a symbol for certain data in a data storage, such as used for “text unit”, for distinguishing the data from other data.

Therefore, it would have been obvious to one of ordinary skill in the art at time the invention was made to modify Sharman by specifically marking a text unit of the processed data, as taught by R1, for the purpose of distinguishing the text unit that is not in the dictionary and preparing for further processing stages, such as processing in a back-up mechanism, generating

Art Unit: 2654

phonemes, coping with prosodic information (Sharman, column 5, lines 25-26, column 5, lines 30-56 and column 5, lines 26). In addition, there must inherently exist some mechanism to distinguish a word that is not in the dictionary from other word that is in the dictionary in Sharman system, because Sharman suggest using a dictionary lookup and some back-up mechanism for handling the two different situation (column 5, lines 23-25).

Regarding claim 16, Sharman and R1 disclose everything claimed, as applied above (see claim 7). Sharman further suggests that: (i) at substring level, it is useful to include some back-up mechanism to be able to process words that are not in the dictionary (column 5, line 24); (ii) at phoneme level, it is again using a dictionary look-up table, augmented with general purpose rules for words not in the dictionary (column 5, line 34); which is equivalent to use “secondary text to speech engine”. Further more, Sharman discloses that the buffer may be used for storing multi-stage input and output (column 7, lines 61-67) for different text units depending on the process stage (column 6, line 61 to column 7, line 22), which inherently includes process stage(s) in secondary TTS engine. This corresponds to the claimed “passing said marked textual unit to a secondary text to speech engine, receiving a speech sample converted from said marked textual unit from said secondary text to speech engine, and appending said converted speech sample to said output signal.”

9. Claims 8 and 17 are rejected under 35 U.S.C. 103(a) as being unpatentable over Sharman in view of R1 and further in view of O'Donnell (“programming for the world--a guide to internationalization”, ISBN 0-13-722190-8).

Regarding **claim 8**, Sharman and R1 disclose everything claimed, as applied above (see claim 7). But, Sharman and R1 fail to disclose that “said marking comprises pre-pending a character to said textual unit.” However, the examiner contends that the concept of marking a text unit by using a pre-pending character was well known, as taught by O'Donnell.

O'Donnell writes a book of “programming for the world”, which discloses that appending a character symbol “\$” to a digit string for distinguishing monetary amount from normal number (page 49, table 2.11).

Therefore, it would have been obvious to one of ordinary skill in the art at time the invention was made to modify Sharman and R1 by specifically marking a text unit of the processed data by adding a character, such as “\$” or the like, in front of the text units, as taught by O'Donnell, for the purpose of easily distinguishing the text units and preparing for further processing.

Regarding **claim 17**, Sharman, R1 and O'Donnell disclose everything claimed, as applied above (see claim 8). Sharman further suggests that: (i) at substring level, it is useful to include some back-up mechanism to be able to process words that are not in the dictionary (column 5, line 24); (ii) at phoneme level, it is again using a dictionary look-up table, augmented with general purpose rules for words not in the dictionary (column 5, line 34); which is equivalent to use “secondary text to speech engine”. Further more, Sharman discloses that the buffer may be used for storing multi-stage input and output (column 7, lines 61-67) for different text units depending on the process stage (column 6, line 61 to column 7, line 22), which inherently includes process stage(s) in secondary TTS engine. This corresponds to the claimed “passing said marked textual unit to a secondary text to speech engine; receiving a speech sample

converted from said marked textual unit from said secondary text to speech engine; and appending said converted speech sample to said output signal."

10. Claims 14-15 and 19-20 are rejected under 35 U.S.C. 103(a) as being unpatentable over Sharman in view of Hata and further in view of Malsheen et al. (USPN 4,979,216) hereinafter referenced as Malsheen.

Regarding **claim 14**, Sharman discloses a text to speech system, comprising obtaining sufficient data and storing output buffer (column 8, lines 21-29), and having different text units for processing speech data in different stages (column 6, 61-66, and column 7, Table 1), wherein text units may inherently include **prefix or suffix**, and **word** (column 5, lines 18-27), which inherently exists data structure and satisfies the claimed "said text unit is one of a word, a prefix, or a suffix". But Sharman does not expressly disclose "a data structure" including several fields. However, this feature is well known in the art as evidenced by Hata who discloses the data structures in the computer memory, including word list stored in the word list buffer and the (speech) sample list data structure stored in the sample list buffer and the relationship between them (column 6, line 64 through column 7, line 32, and Fig. 3). Therefore, it would have been obvious to one of ordinary skill in the art at time the invention was made to combine Sharman and Hata, to expressly provide data structures for processing and storing speech data, as taught by Hata, for the purpose of producing high-quality concatenated synthesized speech (Hata: column 2, line 57).

Further, Sharman in view of Hata does not expressly disclose the data structure having ability for "a frequency of a first portion of the speech sample that exceeds an amplitude

threshold and for a frequency of a last portion of the speech sample that exceeds an amplitude threshold." However, this feature is well known in the art as evidenced by Malsheen who discloses the data structures for phoneme having frequency related information (column 5, line 65 through column 6, line 26) and other data fields (Table 1-4), which suggest that combined system is capable of implementing the claimed data structure. Therefore, it would have been obvious to one of ordinary skill in the art at time the invention was made to combine Sharman and Hata and Malsheen, to provide data structures having field(s) relating frequency information and other data fields for processing and storing speech data, as taught by Malsheen, for the purpose of reducing cost (Malsheen: column 2, line 57).

Furthermore, in another view, a data structure is the template that data can be applied to. For computer and/or microprocessor based devices, data structure is an inherent nature for storing, accessing the required data through associated hardware and/or software functionalities. The claimed data structure has only three fields without any functionality and connection with any software and hardware, so that, in fact, any three data elements can apply to this data structure. Because the data structure can simply be interpreted as three-field template, Sharman and Hata and Malsheen can, either individually or in combination, satisfy the limitation of the claim.

Regarding **claim 15**, Sharman and Hata and Malsheen disclose everything claimed, as applied above (see claim 14). The rejection is based on same or similar reason described in claim 14, because this only adds two more fields which is interpreted as the template with few more fields that any data can be applied to, so that Sharman and Hata and Malsheen can, either individually or in combination, satisfy the claimed limitation(s). In addition, Sharman in view of Hata in view of Malsheen further discloses a phonological feature table (an array type of data

Art Unit: 2654

structure) 52 (Hata: Fig. 3), comprising fields of phonemes that a word may begin and end with (Hata: column 5, lines 14-31, and column 7, lines 55-59), which further corresponds to the claimed "a field for a phoneme that said textual unit starts with, and a field for a phoneme that the textual unit ends with."

Regarding claims 19 and 20, Sharman and Hata and Malsheen disclose everything claimed, as applied above (see claim 14). The rejection is based on same or similar reason described in claim 14, because these claims only add three more fields which is interpreted as the template with few more fields that any data can be applied to, so that Sharman and Hata and Malsheen can, either individually or in combine, satisfy the claimed limitation(s).

Conclusion

11. Any response to this office action should be mailed to:
Commissioner of Patents and Trademarks, P.O. Box 1450, Alexandria, VA22313-1450
or faxed to:

(703)-872-9314

Hand-delivered responses should be brought to:

Crystal Park II, 2121 Crystal Drive, Arlington, VA. Sixth Floor (Receptionist).

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Qi Han whose telephone numbers is (703) 305-5631. The examiner can normally be reached on Monday through Thursday from 8:00 a.m. to 5:30 p.m. and Friday from 8:00 a.m. to 12:00 a.m.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil, can be reached on (703) 305-6954.

Any inquiry of a general nature of relating to the status of this application or proceeding should be directed to the Technology Center 2600 Customer Service Office whose telephone number is (703) 306-0377.

QH/qh

November 24, 2003



RICHEMOND DORVIL
SUPERVISORY PATENT EXAMINER